THREAT ANALYSIS

CHINA

Recorded Future®

By Insikt Group®

June 17, 2025

Lead Author Zoe Haver

Artificial Eyes: Generative AI in China's Military Intelligence

The PLA has demonstrated interest in using generative AI for intelligence work, designed generative AI-based intelligence methods and systems, and likely procured generative AI for intelligence purposes.

The PLA's generative AI-based intelligence tools are broadly intended to improve the speed, efficiency, accuracy, and scale of intelligence tasks while reducing costs.

The PLA and China's defense industry have very likely used proprietary and open-source generative AI models, including models from foreign companies, to develop intelligence tools.

Executive Summary

The People's Liberation Army (PLA) has demonstrated clear interest in using generative artificial intelligence (AI) to support intelligence work, has designed methods and systems that apply generative AI to intelligence tasks, and has likely procured generative AI for intelligence purposes. Both the PLA and China's defense industry have very likely adapted foreign and domestic large language models (LLMs) to develop specialized models that can effectively carry out intelligence tasks. The PLA and China's defense industry have created generative AI-based intelligence tools that can reportedly process and analyze intelligence data, generate intelligence products, answer questions, provide recommendations, facilitate early warning, and support decision-making, among other functions. These tools are broadly intended to improve the speed, efficiency, accuracy, and scale of intelligence tasks while reducing costs. Though elements of the PLA have expressed optimism about the benefits of generative AI and are likely taking initial steps to apply this technology to intelligence work, the PLA has very likely recognized the limitations and risks of this technology. Consequently, the extent to which the PLA will integrate generative AI into intelligence activities — and the ultimate effectiveness of this integration — remains unclear.

The PLA's interest in using generative AI to support military intelligence presents challenges for both the PLA and the West. For the PLA, given the limitations and risks of generative AI, successful adoption of this technology will require experimenting with the intelligence applications of generative AI, accurately assessing the outcomes of these experiments, and appropriately applying generative AI to intelligence work based on these outcomes and assessments; failure to do so could result in inaccurate intelligence that degrades the quality of decision-making. Moreover, if PLA intelligence analysts use generative AI models that were developed to conform with Chinese Communist Party (CCP) ideology or trained on ideologically biased analytical products, the PLA risks reducing the objectivity of intelligence analysis. For the West, the PLA's application of generative AI to intelligence work creates technology transfer challenges and highlights the risk of Chinese counterintelligence organizations using generative AI to generate inauthentic but convincing information to mislead Western intelligence analysts and degrade the intelligence value of open-source information.

Key Findings

- PLA media and researchers affiliated with the PLA have argued that the application of generative AI to military intelligence has a wide range of potential benefits, including improving the collection and analysis of intelligence and providing enhanced decision-making capabilities to military commanders, but have also recognized various challenges and risks associated with using this technology for intelligence work.
- Likely realizing the intelligence limitations of general-purpose generative AI models, the PLA and China's defense industry are likely prioritizing the development and use of specialized models that have been fine-tuned for intelligence tasks.
- The PLA and China's defense industry have very likely used a mix of proprietary and open-source LLMs from foreign and domestic developers to create generative AI-based intelligence tools. Foreign LLMs used this way very likely include models from Meta, OpenAI, and BigScience, among others, while domestic LLMs very likely include models from DeepSeek, Tsinghua University, Zhipu AI, and Alibaba Cloud, among others.
- PLA patent applications reveal that the PLA has designed methods and systems that use generative AI to facilitate intelligence tasks such as generating open-source intelligence (OSINT) products, processing satellite imagery, supporting event extraction, and processing event data.
- In a patent application filed in December 2024, a Chinese state-owned defense industry research institute proposed using OSINT, human intelligence (HUMINT), signals intelligence (SIGINT), geospatial intelligence (GEOINT), and technical intelligence (TECHINT) data to train a military LLM to specialize in intelligence tasks, purportedly enabling the enhanced military LLM to support every phase of the intelligence cycle and improve decision-making during military operations.
- The PLA and China's defense industry have likely procured generative AI technology to support OSINT and science and technology (S&T) intelligence, an indicator that at least some elements of China's military are likely beginning to apply generative AI to intelligence tasks.
- The PLA — which very likely rapidly adopted DeepSeek's generative AI models in early 2025 — is likely using DeepSeek's LLMs for intelligence purposes, based on claims by a Chinese defense contractor that it has provided a DeepSeek-based OSINT model to the PLA.
- The PLA is likely concerned that foreign counterintelligence organizations could use generative AI to produce convincing inauthentic content to mislead Chinese intelligence personnel and degrade the intelligence value of open-source information. Chinese counterintelligence organizations could apply generative AI in a similar manner.

Methodology

To assess the PLA's views on and application of generative AI in intelligence work, Insikt Group collected and analyzed articles published in PLA media, academic research authored by PLA personnel, PLA and Chinese defense industry patent applications, PLA and Chinese defense industry procurement records, information published by Chinese defense contractors, and data available in the Recorded Future® Platform, among other sources.[1] The sources cited in the report do not necessarily represent official PLA policies related to generative AI; rather, they demonstrate how individuals and organizations situated within the PLA and China's defense industry are likely exploring and developing the intelligence applications of generative AI. This report does not assess the veracity of technical claims from the PLA and China's defense industry, but entities that develop or sell intelligence-related generative AI tools likely have an incentive to exaggerate the effectiveness of generative AI and downplay its deficiencies, so information provided by these entities should be viewed with skepticism.

Generative AI is a broad term, and some of its subcategories lack clear boundaries, which occasionally created challenges for this investigation.[2] In some instances, we observed references to generative AI models like LLMs being used for possible non-generative functions; we assessed these references were still relevant to our investigation. Moreover, Chinese sources frequently use the term "large model" (大模型) when discussing AI models rather than more specific terms like "LLM." Not all large models are considered to be generative AI, so we did not automatically assume that every reference to a "large model" was related to generative AI. We only classified a reference to a "large model" as generative AI when other information confirmed the model's generative nature. **Appendix A** provides a glossary of generative AI terminology for readers unfamiliar with this technology.

Views on Generative AI in Military Intelligence

PLA Daily

The PLA's official paper, PLA Daily, has published a handful of articles that directly discuss the intelligence implications of generative AI, which are summarized in **Table 1**.[3] These articles discuss the potential benefits of generative AI for military intelligence, highlighting the supposed ability of this technology to generate intelligence products, predict changes on the battlefield, facilitate intelligence activities during both peacetime and wartime, improve the efficiency of intelligence analysis, and

[1] Most of these sources are from 2023 onward, given the significant attention generative AI has received in recent years.
[2] Helen Toner, "What Are Generative AI, Large Language Models, and Foundation Models?," Center for Security and Emerging Technology, *https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/*.
[3] Liu Kui [刘奎] and Wang Bingbing [王冰冰], "Opportunities and Limitations of Large Models in Operational Command and Control" [大模型在作战指挥中的优势和局限], China Military Online [中国军网], August 29, 2024, *https://web.archive.org/web/20240905185524/http://www.81.cn/yw_208727/16334113.html*; Mao Weihao [毛炜豪], "The Military Applications of Artificial Intelligence From the Perspective of ChatGPT" [从ChatGPT看人工智能的军事应用], *PLA Daily* [解放军报], April 13, 2023, *https://web.archive.org/web/20240613195946/http://www.81.cn/jfjbmap/content/2023-04/13/content_337523.htm*; Hu Xiaofeng [胡晓峰], "ChatGPT, How Should We View" [ChatGPT，我们该怎么看], *PLA Daily* [解放军报], March 21, 2023, *https://web.archive.org/web/20241209173640/http://www.81.cn/jfjbmap/content/2023-03/21/content_335979.htm*; Yang Baosheng [杨宝升], "Be on Alert for Deepfake Technology" [警惕深度伪造技术], *PLA Daily* [解放军报], June 19, 2020, *https://web.archive.org/web/20230222225429/http://www.81.cn/jfjbmap/content/2020-06/19/content_264200.htm*.

provide decision-making support to commanders. These articles also note potential challenges associated with generative AI, suggesting that generative AI models cannot replace up-to-date intelligence and warning that the intelligence agencies of competing countries could exploit deepfakes to interfere with rival agencies.

Date	Summary
August 2024	PLA Daily published an article on large generative AI models in operational command and control, which contends that large models can generate intelligence briefs, extract intelligence main points, and predict changes in the battlefield situation.[4] However, the article also suggests that large models cannot replace up-to-date intelligence information, arguing that these models can function as encyclopedias but can only query information from their training data. As such, large models reportedly depend on up-to-date intelligence information to analyze and judge the battlefield situation, organize and integrate intelligence products, and connect and mine intelligence knowledge.
April 2023	PLA Daily published an article on the military applications of ChatGPT, which describes ChatGPT as having intelligence applications in both peacetime and wartime.[5] In peacetime, ChatGPT can reportedly serve as a virtual assistant to help analysts analyze the massive amount of information available on the internet, improving the efficiency of analysis and mining hidden high-value intelligence. In wartime, ChatGPT can reportedly integrate large amounts of battlefield intelligence into a comprehensive battlefield situation report automatically, reducing the workload of intelligence personnel and improving the intelligence analysis and planning capabilities of combat personnel.
March 2023	PLA Daily published an article on ChatGPT that predicts combatants will have strong intelligence collection capabilities and near-real-time information perception capabilities in "future informatized and intelligentized wars."[6] To this point, the article notes that ChatGPT could be used for basic work such as data analysis, decision-making support, and natural language processing, qualitatively improving commanders' decision-making capabilities through mass battlefield information processing.
June 2020	The PLA Daily published an article on the dangers of deepfake technology that warns deepfakes could be used to interfere with intelligence work.[7] The article warns that the intelligence agencies of competing countries could use deepfakes to interfere with rival agencies and set limits on the scope of their operations.

***Table 1**: Summary of PLA Daily articles that discuss the intelligence implications of generative AI (Source: PLA Daily; Insikt Group)*

[4] Kui and Wang, "Opportunities and Limitations of Large Models in Operational Command and Control."
[5] Mao, "The Military Applications of Artificial Intelligence From the Perspective of ChatGPT."
[6] Hu, "ChatGPT, How Should We View."
[7] Yang, "Be on Alert for Deepfake Technology."

Academic Research

PLA researchers — especially personnel affiliated with the Academy of Military Science (AMS; 军事科学院) Military Science Information Research Center (MSIRC; 军事科学信息研究中心) — have expressed optimism about the intelligence applications of generative AI but have also recognized the fallibility of this technology, raising a variety of issues associated with applying generative AI to intelligence work. In one notable study, AMS MSIRC researchers assessed the opportunities and challenges of using generative AI to support national defense S&T intelligence, providing recommendations related to taking advantage of generative AI's opportunities while mitigating its challenges. PLA researchers have also analyzed efforts within the United States (US) military to apply generative AI to intelligence tasks, likely aiming to learn from the US military's experience and adapt best practices. Researchers not affiliated with the PLA but associated with other elements of China's party-state system have likewise published on the intelligence implications of generative AI, reflecting sentiments expressed by PLA personnel and providing insight into debates likely occurring within China's party-state system.

General Views in the PLA

PLA researchers have expressed interest in the intelligence applications of generative AI, with some describing it as a potentially transformative technology.[8] For example, in August 2024, researchers affiliated with the AMS MSIRC published an article about the effects of AI on intelligence research that contends the development of technologies like machine learning, deep learning, and generative AI has created unprecedented opportunities for intelligence research.[9] In June 2024, researchers affiliated with AMS MSIRC, the AMS National Innovation Institute of Defense Technology (国防科技创新研究院), and two civilian universities published a study that details their use of Meta's Llama 13B model to develop an LLM that specializes in military OSINT, arguing that LLMs can facilitate comprehensive and accurate intelligence support for military commanders and claiming that their LLM could support intelligence analysis, strategic planning, simulation training, and command decision-making.[10] Similarly, in February 2024, researchers affiliated with AMS MSIRC published an article that suggests applying

[8] Tang Shanhong [汤珊红] et al., "Generative Artificial Intelligence Empowering National Defense Science and Technology Intelligence" [生成式人工智能赋能国防科技情报], *Information Studies: Theory & Application* [情报理论与实践] 46, no. 11 (2023); Geng Guotong [耿国桐] et al., "The Transformation and Development of Artificial Intelligence Empowering Intelligence Research"[人工智能赋能情报研究的变革与发展], *Advances in Information Science* [情报学进展] 1 (2024); Sun Yazhou [孙亚洲] et al., "Intelligence Improvements and Intelligence Limitations of ChatGPT-Like Empowered Command and Control - - Based on Information Flow Model" (类ChatGPT赋能指挥控制的情报提升和情报局限——基于信息流动模型), *Journal of Intelligence* [情报杂志] 43, no. 6 (2024); Liu Huashuai [刘华帅] et al., "Information and Communications Equipment Application Problem Research Based on Generative AI Models" (基于生成式AI模型的信息通信装备运用问题研究), 11th China Command and Control Conference Proceedings [第十一届中国指挥控制大会论文集] (2023); Hu Yiting [胡艺婷], "Prospect Analysis of Generative AI Applications in United States Military 'Decapitation Operation' Intelligence Support" [美军"斩首行动"情报支援中生成式AI应用前景分析], 12th China Command and Control Conference Proceedings [第十二届中国指挥控制大会论文集] (2024); Chen Ting [陈婷], "Obtaining Competitive Advantage: The Military Application of United States Generative Artificial Intelligence" [获取竞争优势：美国生成式人工智能的军事应用], *Contemporary World* [当代世界] 5 (2024); Zhang Huaping [张华平] et al., "Large Language Model-Driven Open-Source Intelligence Cognition" [大语言模型驱动的开源情报认知], *National Defense Technology* [国防科技] 45, no. 3 (2024).

[9] Geng et al., "The Transformation and Development of Artificial Intelligence."

[10] Zhang et al., "Large Language Model-Driven Open-Source Intelligence Cognition"; James Pomfret and Jessie Pang, "Exclusive: Chinese researchers develop AI model for military use on back of Meta's Llama," *Reuters*, November 1, 2024, *https://www.reuters.com/technology/artificial-intelligence/chinese-researchers-develop-ai-model-military-use-back-metas-llama-2024-11-01/*; Sunny Cheung, "PRC Adapts Meta's Llama for Military and Security AI Applications," *China Brief* 24, no. 21 (2024), *https://jamestown.org/program/prcs-adaptation-of-open-source-llm-for-military-and-security-purposes/*.

ChatGPT-like technologies to command and control systems could result in major improvements to intelligence capabilities.[11]

Despite recognizing possible applications of generative AI for artificial intelligence, PLA researchers have also discussed challenges associated with this technology.[12] In the aforementioned February 2024 article, AMS MSIRC researchers warn that the intelligence limits of ChatGPT-like technologies integrated into command and control systems could result in catastrophic consequences.[13] Likewise, in the June 2024 article, AMS and civilian university personnel contend that current LLMs have serious hallucination issues and are unsuitable for direct use in OSINT.[14] In September 2024, researchers affiliated with the PLA National University of Defense Technology's (NUDT; 中国人民解放军国防科技大学) College of Information and Communications (信息通信学院) and College of Intelligent Science (智能科学学院) published an article about the the relationship between AI and OSINT that argues deepfakes produced with generative AI create significant challenges for efforts to exploit the massive amount of information available on the internet for intelligence purposes.[15]

National Defense S&T Intelligence Example

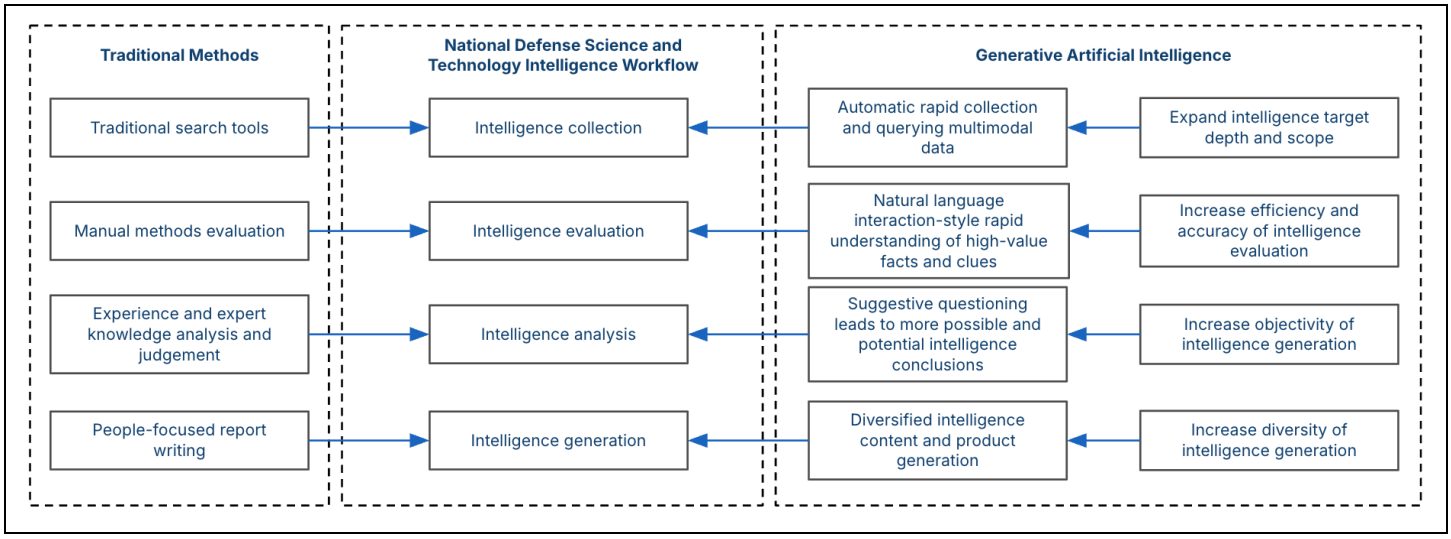

***Figure 1**: Translated table of potential national defense S&T intelligence generative AI applications and benefits (Source: "Empowering National Defense Science and Technology Intelligence"[16]; Insikt Group)*

[11] Sun et al., "Intelligence Improvements and Intelligence Limitations."
[12] Tang et al., "Empowering National Defense Science and Technology Intelligence"; Sun et al., "Intelligence Improvements and Intelligence Limitations"; Liu et al., "Information and Communications Equipment Application Problem Research"; Hu, "Prospect Analysis of Generative AI Applications"; Chen, "Obtaining Competitive Advantage"; Zhang et al., "Large Language Model-Driven Open-Source Intelligence Cognition"; Yuan Weilin [袁唯淋] et al., "Intelligent Intelligence Fusion Summary: An Analysis of Open-Source Intelligence Fusion from an Adversarial Perspective" [智能情报融合综述:对抗视角下的开源情报融合分析], *Chinese Journal of Intelligent Science Technology* [智能科学与技术学报] 3 (2024).
[13] Sun et al., "Intelligence Improvements and Intelligence Limitations."
[14] Zhang et al., "Large Language Model-Driven Open-Source Intelligence Cognition."
[15] Yuan et al., "An Analysis of Open-Source Intelligence Fusion from an Adversarial Perspective."
[16] Tang et al., "Empowering National Defense Science and Technology Intelligence."

In November 2023, researchers affiliated with AMS MSIRC published an article about the potential opportunities and challenges of using generative AI for national defense S&T intelligence.[17] According to the article, traditional approaches to intelligence can no longer identify potential threats and opportunities in a timely and accurate manner, a common view that has almost certainly driven the adoption of new intelligence collection, processing, and analysis technologies within the PLA and China's defense industry.[18] The authors contend that generative AI is profoundly affecting the creation and application of knowledge, including national defense S&T intelligence. They claim that generative AI has the potential to improve intelligence collection, evaluation, analysis, and generation (see **Figure 1**), detailing how generative AI could assist intelligence personnel rather than asserting that this technology could replace intelligence personnel. For example, the authors argue that generative AI could:

- Help intelligence personnel familiarize themselves with unfamiliar technical fields when assigned new intelligence tasks, thereby improving the depth and scope of analysts' understanding of intelligence targets
- Aid intelligence personnel in refining and discovering relationships within large amounts of data and assist personnel with tasks like comparison, inference, providing examples, and induction
- Generate diverse scenarios and assumptions to help intelligence analysts broaden their thinking and avoid cognitive biases during intelligence analysis, as well as evaluate analysts' analytical conclusions to identify potential biases
- Automatically recommend relevant images and audiovisual materials, generate statistical reports, and automatically synthesize relevant materials to facilitate diversified expression of intelligence analysis content beyond traditional intelligence reports

However, they also warn that generative AI could bring unprecedented uncontrollability, uncertainty, and high levels of risk to national defense S&T intelligence. They identify specific challenges such as:

- Heightened counterintelligence risks, including the risk of technological competitors using generative AI to create fake technical documents and deepfakes to mislead national defense S&T intelligence efforts
- A lack of LLMs specifically designed for national defense S&T intelligence
- Generative AI's limited ability to deal with uncertain and biased information
- Insufficient national defense S&T intelligence corpora for training LLMs, including difficulties associated with using state secrets and sensitive intelligence to train LLMs
- Reliability issues like "data leakage, algorithmic black boxes, value bias, and unexplainability"

Describing generative AI as a double-edged sword, the AMS researchers suggest that the field of national defense S&T intelligence should pursue measures to simultaneously take advantage of the opportunities of generative AI and mitigate its challenges. These include:

[17] Tang et al., "Empowering National Defense Science and Technology Intelligence."
[18] Zoe Haver, "Private Eyes: China's Embrace of Open-Source Military Intelligence" (Recorded Future, 2023), *https://www.recordedfuture.com/research/private-eyes-chinas-embrace-open-source-military-intelligence*.

- Gradually introducing generative AI into national defense S&T intelligence work and evaluating the effectiveness of this technology after its introduction
- Working to improve relevant corpora and LLMs
- Iteratively combining intelligence workflows that involve both human and generative AI inputs to ensure reliable and credible results
- Developing technologies to ensure the traceability and verify the reliability of AI-generated content

Interest in US Military Applications

PLA researchers have assessed how the US military is applying generative AI to intelligence tasks,[19] likely aiming to learn from the US military's experience in a manner resembling China's previous efforts to study OSINT practices and AI military applications in the US.[20] For example, in May 2024, a researcher affiliated with the AMS War Research Institute (战争研究院) published an article analyzing how the US is exploring the military applications of generative AI.[21] The author discusses organizational changes, policy guidance, experimental testing, and security measures the US has reportedly implemented; particular military applications the US has reportedly conceptualized, tested, or put into use; and challenges the US has reportedly encountered. Notably, the author claims that the Defense Innovation Unit (DIU) of the US Department of Defense (DoD) initiated a technology program in May 2023 to explore the applications of generative AI in OSINT collection and analysis. This technology reportedly facilitated automatic data mining and evaluation, and visualized the battlefield information environment for commanders. The author highlights that DIU required this technology to help analysts edit and disseminate content, comply with DIU's AI standards, and be usable within the DoD's information environment.

Non-PLA Perspectives

Beyond the PLA, specialists situated within other segments of China's party-state system have also examined the opportunities and challenges of using generative AI for intelligence work. Though these discussions do not focus on military intelligence, they can still provide insight into how the broader party-state system is likely grappling with the intelligence implications of this technology. For example, in November 2024, researchers affiliated with the People's Public Security University of China (PPUSC) School of State Security (中国人民公安大学国家安全学院) published an article that highlights the risk of "false information pollution" — online disinformation created by generative AI — disrupting OSINT work.[22] Moreover, in a June 2024 article that focuses on the potential effect of ChatGPT and other similar AI tools on intelligence work, researchers affiliated with the PPUSC Public Security Intelligence Research Center (中国人民公安大学公安情报研究中心) argue that careless use of ChatGPT could undermine the

[19] Hu, "Generative AI Applications in United States Military"; Chen, "Obtaining Competitive Advantage"; Tang et al., "Empowering National Defense Science and Technology Intelligence."
[20] Haver, "Private Eyes"; Elsa B. Kania, "Battlefield Singularity Artificial Intelligence, Military Revolution, and China's Future Military Power" (Center for a New American Security, 2017), *https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power*.
[21] Chen, "Obtaining Competitive Advantage."
[22] Li Junchen [李俊辰] and Wu Shaozhong [吴绍忠], "The Impact of False Information on Open-Source Intelligence Work From the Perspective of GAI Technology" [GAI技术视角下的虚假信息对开源情报工作的影响], *Information Research* [情报探索] 11 (2024).

CCP's ideological leadership of intelligence work.[23] The researchers remind readers that ChatGPT was developed by a US company and was largely trained with an English-language corpus. They warn that ChatGPT could be affected by "Western capitalist values," generate false but convincing information based on "neoliberalism" and "ideological neutrality," erode ideological discourse and management, and facilitate the invisible penetration of ideology into intelligence work. The views expressed in this article coincide with Chinese regulators' broader efforts to ensure the correct ideological alignment of generative AI models.[24]

Applications of Generative AI in Military Intelligence

Patent Activity

The PLA and China's defense industry have filed multiple patent applications that involve applying generative AI to intelligence tasks, an indicator that the PLA is likely serious about testing and using this technology to support intelligence work. These include a particularly notable patent application from the Ordnance Science and Research Academy of China (OSRAC; 中国兵器科学研究院), as well as patent applications from AMS MSIRC, NUDT, the AMS Systems Engineering Institute (系统工程研究院), and the AMS War Research Institute.

Military Intelligence LLM Developed by OSRAC

In October 2024, OSRAC filed a patent application, which was granted in May 2025, that proposes using data from OSINT, HUMINT, SIGINT, GEOINT, and TECHINT sources to help train a "military LLM" (军用大语言模型) to specialize in intelligence tasks, purportedly enabling the military LLM to support every phase of the intelligence cycle and improve decision-making during military operations.[25] OSRAC is part of China Ordnance Industry Group Corporation Limited (中国兵器工业集团有限公司; NORINCO Group), a major state-owned defense enterprise in China.[26] Due to the length and complexity of OSRAC's patent, this section translates and summarizes key elements of the document but does not discuss every aspect of the invention. **Figure 2** provides a translation of a technical architecture diagram that offers a general overview of OSRAC's invention.

[23] Zhang Guoqing [张国庆] et al., "The Impact of Generative Artificial Intelligence for Intelligence Work Under the TCPED Method: The Case of ChatGPT" [TCPED路径下生成式人工智能对情报工作的影响 ——以ChatGPT为例], *Journal of Intelligence* [情报杂志] 43, no. 6 (2024).
[24] Helen Toner et al., "How will China's Generative AI Regulations Shape the Future? A DigiChina Forum," *DigiChina*, April 19, 2023, *https://digichina.stanford.edu/work/how-will-chinas-generative-ai-regulations-shape-the-future-a-digichina-forum/*; Wendy Chang, "China keeps generative AI on simmer," Mercator Institute for China Studies, April 17, 2024, *https://merics.org/en/comment/china-keeps-generative-ai-simmer*; Matt Sheehan, "China's AI Regulations and How They Get Made," Carnegie Endowment for International Peace, July 10, 2023, *https://carnegieendowment.org/research/2023/07/chinas-ai-regulations-and-how-they-get-made?lang=en*; Ryan McMorrow and Tina Hu, "China deploys censors to create socialist AI," *Financial Times*, July 17, 2024, *https://www.ft.com/content/10975044-f194-4513-857b-e17491d2a9e9*.
[25] Zhao Yuechen [赵悦辰] et al., "A method and system for understanding key intelligence requirements and intentions based on a large model," Google Patents, October 24, 2024, *https://patents.google.com/patent/CN119475217A/en*.
[26] "Ordnance Science and Research Academy of China 2024 Campus Recruitment Notice" [中国兵器科学研究院2024年校园招聘公告], State-owned Assets Supervision and Administration Commission of the State Council [国务院国有资产监督管理委员会], September 22, 2023, *https://web.archive.org/web/20250415204147/http://www.sasac.gov.cn/n2588035/n2588325/n2588350/c28930344/content.html*.

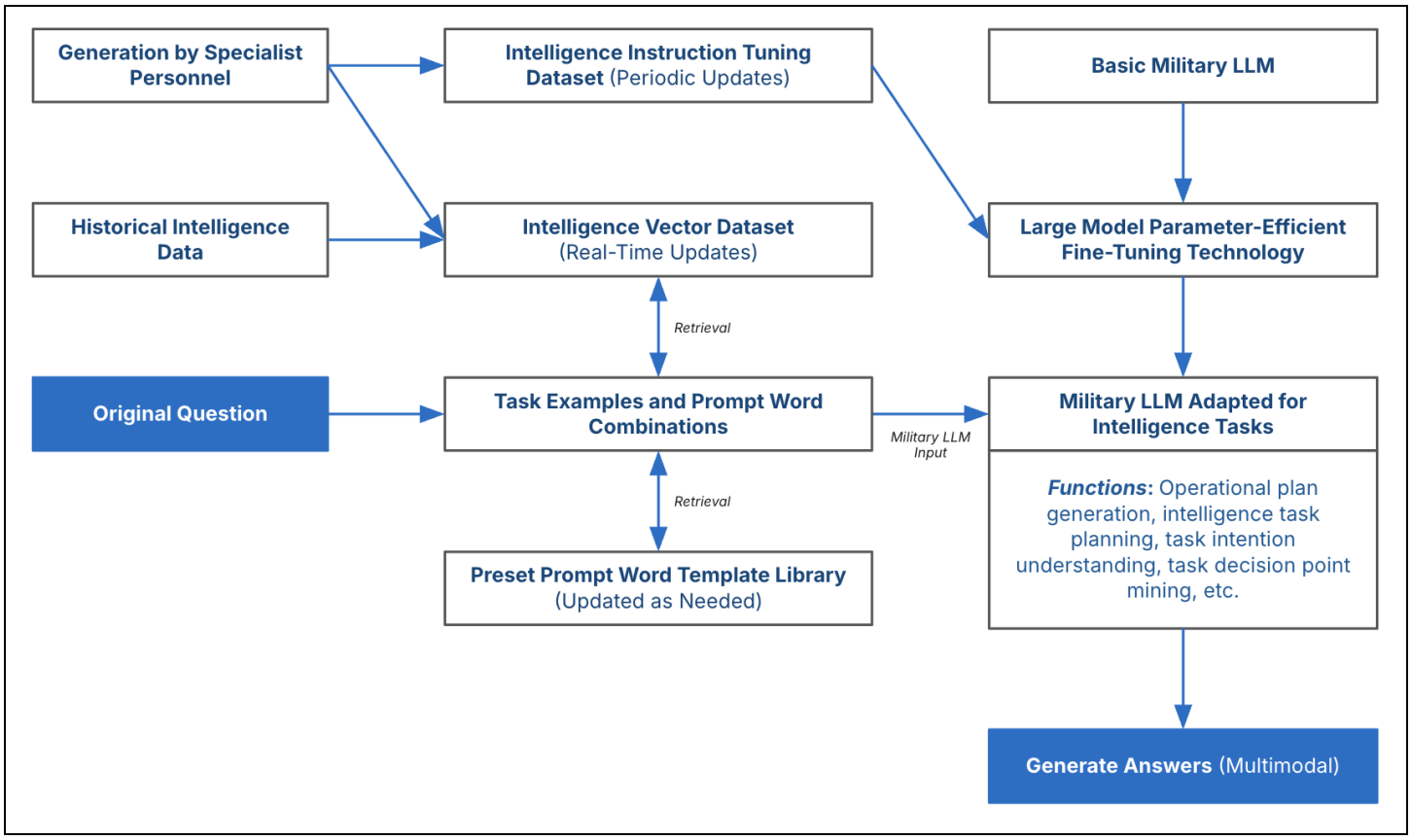

***Figure 2**: Translated technical architecture diagram (Source: OSRAC patent[27]; Insikt Group)*

OSRAC's invention focuses on the development and application of a military LLM that specializes in intelligence tasks. The creation of this intelligence-oriented military LLM requires fine-tuning a military LLM based on ChatGLM, a family of open-source models developed by Zhipu AI and Tsinghua University.[28] According to the patent, the fine-tuning process involves:

1. Collecting intelligence data from OSINT, HUMINT, SIGINT, GEOINT, and TECHINT sources and processing this data, thereby providing a high-quality and high-relevancy data foundation for subsequent model training and decision-making support, increasing the accuracy and reliability of intelligence, and strengthening "understanding capabilities" for complex combat environments

2. Building data models tailored for OSINT, HUMINT, SIGINT, GEOINT, and TECHINT to facilitate the management and analysis of complex intelligence data, support the accuracy and depth of intelligence work, and provide a comprehensive analytical perspective (see **Table 2**)

3. Using these data models to create nine datasets, including a "multimodal intelligence understanding dataset," an "intelligence report dataset," a "combat intelligence requirements

[27] Zhao et al., "A method and system for understanding key intelligence requirements and intentions based on a large model."
[28] Team GLM, "ChatGLM: A Family of Large Language Models from GLM-130B to GLM-4 All Tools," arXiv, June 18, 2024, *https://web.archive.org/web/20240619020417/https://arxiv.org/html/2406.12793v1*.

understanding dataset," a "command, control, and communications dataset," and a "tactics and strategy guidelines dataset"

4. Using the nine initial datasets to develop an "intelligence vector dataset," which is updated in real-time with the latest intelligence and used to help generate inputs for the military LLM, and an "expected intelligence instruction tuning dataset," which is filtered, checked, and supplemented by military specialists, periodically updated, and used to help fine-tune the military LLM

5. Applying "multimodal instruction tuning" and "parameter-efficient fine-tuning" to fine-tune the military LLM to adapt it to intelligence tasks

Data Type	Data Models
OSINT	Entity-Relationship Model (实体-关系模型) • Used to identify and associate entities on the internet, such as individuals, organizations, events, and locations • Relationship mapping, such as connections on social media and appearances together in news reporting Text Analysis Model (文本分析模型) • Used to analyze and structure news reports, social media posts, and other text data • Apply natural language processing technology to extract keywords, themes, and emotions
HUMINT	Graph Database Model (图数据库模型) • Map interpersonal networks, such as the relationships between spies, intelligence agents, and contacts • Can be used to track information flows and for social network analysis Event-Driven Model (事件驱动模型) • Record and analyze specific events during HUMINT activities, such as meetings, communications, and reports
SIGINT	Time Series Model (时间序列模型) • Used to process and analyze changes over time in electronic signals and communications data • Applicable for capturing and analyzing signal patterns and communication behaviors Network Traffic Analysis Model (网络流量分析模型) • Analyze internet traffic and communication patterns, such as packets, sources, and destinations

GEOINT	Geographic Information System Model (地理信息系统模型) ● Store and process geographic location data, such as maps and satellite imagery ● Support spatial data analysis, such as geocoding and regional monitoring Multidimensional Data Model (多维数据模型) ● Applicable for analyzing the relationships between geographic data and other dimensions, such as times and events
TECHINT	Characteristics Library Model (特征库模型) ● Build a detailed characteristics library for weapons systems and equipment, including performance parameters and functional descriptions ● Can be used to identify and compare technical equipment Association Rule Learning Model (关联规则学习模型) ● Analyze the connections between weapons systems, such as common technical sources and use patterns

***Table 2**: Translation of a table that describes data models for OSINT, HUMINT, SIGINT, GEOINT, and TECHINT (Source: OSRAC patent[29]; Insikt Group)*

OSRAC's invention combines its intelligence-oriented military LLM with prompt engineering, creating prompt templates designed to support various intelligence tasks. Each prompt template includes five elements: role, instruction, context, input data, and output indicator. The patent identifies numerous intelligence tasks and provides corresponding prompt templates. These tasks include intelligence retrieval, processing, analysis, and application. Translations of these four prompt templates are provided in **Table 3**. Examples of role-instruction combinations from other prompt templates include helping battlefield intelligence analysts to identify intelligence requirements related to enemy and friendly forces, helping military operations commanders responsible for real-time decision-making to develop follow-up operational plans for possible battlefield outcomes, and helping tactical analysts to assess the current battlefield environment and provide tactical recommendations.

[29] Zhao et al., "A method and system for understanding key intelligence requirements and intentions based on a large model."

Intelligence Task	Prompt Elements
Intelligence Information Retrieval	**Role:** Intelligence collector, OSINT analyst, or cyber intelligence specialist **Instruction:** Identify and extract information related to specific events, regions, people, or topics, including text, image, audio, and social media content **Context:** International relations trends, specific conflicts or crises, technology trends, and cybersecurity threats **Input Data:** Time ranges, geographic locations, specific languages or dialects, social media trends, satellite imagery, and news reports **Output Indicator:** Comprehensive intelligence reports, including data sources, timestamps, geographic tags, emotion analysis, and key event summaries
Intelligence Information Processing	**Role:** Data scientist or intelligence integration analyst **Instruction:** Extract information from various formats and sources, unify data formats, and identify key variables and relationships **Context:** Data availability, source credibility, and the compatibility of multiple data types **Input Data:** Original datasets, intelligence reports, OSINT, social media content, and geospatial information **Output Indicator:** Standardized and cleaned datasets, including metadata, relationship graphs, and time series analysis
Intelligence Information Analysis	**Role:** Strategic analyst, cyber analysis specialist, or GEOINT specialist **Instruction:** Carry out deep analysis and identification of hidden patterns, potential threats, and opportunities **Context:** Current political and economic environment, S&T development, social trends, and historical comparative analysis **Input Data:** Processed datasets, historical intelligence records, and real-time monitoring data **Output Indicator:** Deep analysis reports, including predictive models, risk assessments, and scenario analysis

Intelligence Information Application	**Role:** Policymaker, military commander, or security advisor **Instruction:** Formulate response strategies, plans, and actions based on intelligence analysis results **Context:** Current strategic objectives, available resources and limitations, and potential points of cooperation or conflict **Input Data:** Intelligence analysis reports, early warning system outputs, and geopolitical analysis **Output Indicator:** Action plans, policy recommendations, emergency response measures, and long-term strategic planning

***Table 3**: Translation of four prompt templates (Source: OSRAC patent[30]; Insikt Group)*

The patent suggests that OSRAC's enhanced military LLM has the potential to support every step of the intelligence cycle, namely, intelligence planning, collection, processing, analysis, dissemination, and evaluation.[31] **Table 4** summarizes OSRAC's description of how the LLM can support the intelligence cycle. For clarity, this table uses the standard name for each step of the intelligence cycle, even though OSRAC uses slightly different names for most of the steps in its patent.

Planning
• Help intelligence personnel refine and research new subjects and understand the battlefield environment and enemy threats • Quickly identify information related to potential enemy operations, improving the accuracy and relevancy of intelligence collection • Establish a foundation for subsequent intelligence collection and analysis
Collection
• Analyze and filter large volumes of collected information • Process multilingual content and translate, summarize, and classify important information, improving the usability and understandability of intelligence • Identify key time nodes, geographic region positions, and enemy operation types, making intelligence collection more specific and purposeful

[30] Ibid.
[31] For an overview of the intelligence cycle, see: "How the IC Works," Intel[.]gov, *https://www.intelligence.gov/how-the-ic-works*.

Processing
• Rapidly analyze and integrate intelligence data collected from various sensors and sources • Identify important information gaps, and propose possible action plans and strategic recommendations • Identify latent threats and opportunities, providing support for strategic decision-making
Analysis
• In joint combat environments, generate intelligence products such as combat environment analyses and enemy capabilities evaluation reports • Generate customized intelligence products based on user requirements and commanders' strategic intentions • Provide necessary reference materials such as historical data and early research
Dissemination
• Ensure the rapid and effective distribution and fusion of intelligence • Integrate intelligence into decision-making support systems • Help organizations understand and use intelligence, ensuring the consistency and accuracy of intelligence
Evaluation
• Evaluate the validity and accuracy of intelligence • Help analyze the results of intelligence collection and processing, ensuring these processes meet the requirements of commanders and joint forces

***Table 4**: Summary of how OSRAC claims its LLM can support each step of the intelligence cycle (Source: OSRAC patent[32]; Insikt Group)*

PLA Patents

PLA entities have filed multiple patent applications that focus on or otherwise involve the application of generative AI to intelligence tasks, summarized in **Table 5**. These patent applications were filed by AMS MSIRC, NUDT, the AMS Systems Engineering Institute, and the AMS War Research Institute. The patent applications propose using generative AI to generate OSINT products, process satellite imagery, support event extraction, and process event data. These documents reference using LLMs from multiple foreign companies, including Meta, OpenAI, Google DeepMind, and Anthropic, as well as BigScience's BLOOM.

[32] Zhao et al., "A method and system for understanding key intelligence requirements and intentions based on a large model."

Date	Patent Application Details
December 2024	AMS MSIRC filed a patent application that proposes fine-tuning a general LLM to create a specialized model that can generate accurate Chinese-language news briefs based on English-language websites for OSINT purposes.[33] The document contends that general pre-trained LLMs lack the depth of understanding and judgment capabilities required for military OSINT. The patent application references using various LLMs, including foreign open-source models like Meta's Llama and BigScience's BLOOM. It also discusses using domestic models like Tsinghua University and Zhipu Ai's ChatGLM, Alibaba Cloud's Qwen, and a Baichuan model.
October 2024	NUDT filed a patent application that proposes using LLMs to process satellite imagery.[34] The patent application includes an example, comprising both a diagram and accompanying text, of a user asking an agent to process satellite data with the goal of knowing information related to survey data, information related to intelligence data, and collected reconnaissance target information. In response to this request, the agent is designed to determine that this task involves satellite survey data and imagery data, carry out data pre-processing, and analyze the target area. The patent application repeatedly references using LLMs developed by Meta, including Llama 2 7B and Code Llama 7B. The document repeatedly uses OpenAI's logo in diagrams, though it is unclear whether NUDT used any of OpenAI's tools in this invention.
September 2024	The AMS Systems Engineering Institute filed a patent application that proposes using LLMs alongside lightweight language models to generate intelligence summaries for OSINT.[35] The patent application claims that the proposed system can improve the processing speed and efficiency of intelligence summaries, produce high-quality summaries at a low cost, and provide efficient intelligence analysis for various application scenarios. The document references using multiple LLMs and lightweight language models, including foreign models like OpenAI's "ChatGPT and GPT-4"; Meta's Llama, Llama 2, and Llama 7B; Google DeepMind's Gopher, Chinchilla, and Sparrow; and Anthropic's Claude. It also references domestic models such as Tsinghua University and Zhipu Ai's ChatGLM-6B, Baidu's ERNIE Bot and Ernie 3.0 Titan, and Alibaba Cloud's Qwen-VL.

[33] Geng Guotong [耿国桐] et al., "Chinese brief information generation method, device and computer program product of English website," Google Patents, December 13, 2024, *https://patents.google.com/patent/CN119740555A/en*.
[34] Gu Xueqiang [谷学强] et al., "Automatic business process generation method, device and equipment based on lightweight LLM," Google Patents, October 14, 2024, *https://patents.google.com/patent/CN119311257A/en*.
[35] Zhang Hongguang [张洪广] et al., "Information abstracting method and system based on large language model distillation," Google Patents, September 27, 2024, *https://patents.google.com/patent/CN119397008A/en*.

September 2024	NUDT filed a patent application (which was granted in January 2025) that proposes using LLMs to support event extraction, which the document defines as a subcategory of information extraction.[36] NUDT's patent suggests that information extraction facilitates the rapid acquisition of information of interest, reduces the time and costs of manual processing, increases the efficiency of information use, and can provide comprehensive and accurate data support to decision-makers. The document notes that event extraction can be applied in various fields, including intelligence analysis.
June 2024	The AMS War Research Institute filed a patent application that proposes using an LLM to process political, economic, and other event data collected from the internet to create event datasets in support of strategic operations analysis.[37] The document describes high-quality structured data (extracted from massive amounts of unstructured internet data) as a beneficial supplement to traditional intelligence that can aid strategic research and decision-making. It argues that using an LLM can facilitate the accurate annotation of text data and the rapid, low-labor cost production of event datasets. The patent application discusses using OpenAI's "GPT-3.5 Turbo interface" (gpt-3.5turbo接口) to fine-tune the LLM.

***Table 5**: PLA and Chinese defense industry patent applications related to applying generative AI to intelligence work (Source: Google Patents; Insikt Group)*

Procurement Activity

Insikt Group observed hundreds of likely references to generative AI in PLA and Chinese defense industry procurement records, largely published from late 2023 onward. Among these, we identified three likely instances of the PLA and China's defense industry acquiring or seeking to acquire generative AI technology for intelligence purposes, namely OSINT and S&T intelligence (see **Table 6**).[38] In the first instance, a PLA or Chinese defense industry entity indicated its interest in obtaining a large model to process and analyze internet data, facilitate question answering functions, and generate reports. In the second instance, a likely PLA entity appears to have procured a generative AI-based early warning system capable of monitoring internet data for signs of infectious disease events. In the third instance, the China State Shipbuilding Corporation (CSSC) 714th Research Institute (中国船舶集团有限公司第七一四研究所), which has a track record of engaging in very likely OSINT activity,[39] pursued the development of generative AI-related tools intended to support S&T intelligence, including a tool likely based on generative AI technology from OpenAI. The US government placed export restrictions on the CSSC 714th Research Institute in December 2020 for acquiring and attempting to acquire US technology in support of the PLA.[40]

[36] Sun Yi [孙毅] et al., "Method and device for automatic induction of bidirectional event patterns based on large language model," Google Patents, September 25, 2024, *https://patents.google.com/patent/CN118861304A/en*.
[37] Jia Jun [贾珺], Cheng Quanlin [陈泉林], and Peng Chao [彭超], "A method for constructing event dataset for strategic operations analysis based on large models," Google Patents, June 6, 2024, *https://patents.google.com/patent/CN118760894A/en*.
[38] Source documents held by Insikt Group.
[39] Haver, "Private Eyes"; source documents held by Insikt Group.
[40] "Addition of Entities to the Entity List, Revision of Entry on the Entity List, and Removal of Entities From the Entity List," Federal Register, December 22, 2020,
https://www.federalregister.gov/documents/2020/12/22/2020-28031/addition-of-entities-to-the-entity-list-revision-of-entry-on-the-entity-list-an

Project Name	Project Details
Internet Data Collection and Analysis Services (互联网数据采集及分析服务)	In April 2025, an unspecified PLA or Chinese defense industry entity announced it wanted to acquire "internet data collection and analysis services" that involved the use of a large model for data processing and analysis, question answering, and report generation. According to the document, this project involved the following services: internet data collection and maintenance, text manual precise translation, video translation and production, online real-time machine translation, multilingual optical character recognition processing, large model intelligent question answering and report writing, and data export and storage. The project also included collecting, cleaning, and translating information from multiple sources, including news websites and social media, and using a large model to process and analyze this data.
Generative Artificial Intelligence Large Model Online Monitoring and Early Warning System (生成式人工智能大模型的在线监测预警系统)	In November 2024, a likely PLA entity published a tender for a generative AI-based system that could use network data to detect "early signs and related signals" of infectious disease, complement traditional infectious disease monitoring, and improve early warning capabilities for infectious disease events. The document states that the system should include a generative AI large model, an infectious disease network information module based on a large model, a data fusion module based on a large model, and an early warning module based on a large model, among other components. The PLA Center for Disease Prevention and Control (中国人民解放军疾病预防控制中心) very likely supervised this project, though it is unclear who issued the tender. According to bidding results published in December 2024, Zhongke Wenge (北京中科闻歌科技股份有限公司; Beijing Zhongke Wenge Science and Technology Co. Ltd.) came in first place, and Beijing Tianguang Huitong Science and Technology Co. Ltd. (北京天广汇通科技有限公司) came in second.

d-removal-of-entities; "China State Shipbuilding Corporation," China Defense Universities Tracker, May 5, 2021, *https://unitracker.aspi.org.au/universities/china-state-shipbuilding-corporation/*.

[Unspecified] Science and Technology Intelligence Generative Artificial Intelligence Source Tracing Evidence-Based Model Corpus Processing Tool Development and Services (XX科技情报生成式人工智能溯源循证模型语料加工工具开发与服务)	In August 2024, the CSSC 714th Research Institute published a tender related to developing generative AI-related tools to support S&T intelligence. The project involved the development of tools for document processing, corpus processing, structured data processing, and large model corpus processing. The institute requested functions such as processing document formats like PDF, Word, PowerPoint, Excel, and HTML at a speed of at least 1,000 characters per second; automatically extracting pictures and tables; filtering out sensitive content; and exporting processed data at a speed of at least 1,000 lines per second. The document also specifies that the large model corpus processing tool should "support real-time large model answer feedback, scoring, manual sorting, and evaluation functions based on an OpenAI interface [OpenAI接口]," which suggests that the CSSC 714th Research Institute was likely seeking to use generative AI technology from OpenAI.

***Table 6**: Likely instances of the PLA and China's defense industry acquiring generative AI technology for intelligence purposes (Source: PLA and Chinese defense industry procurement records[41])*

In addition to these three likely instances of the PLA and China's defense industry acquiring generative AI for intelligence purposes, Insikt Group also identified fifteen plausible but unconfirmed instances of such activity.[42] These fifteen PLA or Chinese defense industry projects generally involved the development or application of large models for data collection, processing, or analysis with the intention of supporting situational awareness, decision-making, S&T development, battlefield information analysis, simulations, or early warning, among other goals. However, the procurement documents associated with these projects provide insufficient information about their respective large models, data, or applications, thereby preventing a more conclusive assessment.

Procurement records reveal that the PLA very likely rapidly adopted DeepSeek's LLMs in early 2025, but these documents do not reveal whether the PLA is using this technology for intelligence purposes. We observed over 150 references to DeepSeek in PLA and Chinese defense industry procurement documents as of late May 2025 (these references include multiple references to the same documents, so the actual number of procurement documents is lower).[43] The earliest of these documents was published in February 2025, and the majority of the documents were published in March, April, and May 2025. The PLA very likely embraced DeepSeek following the company's release of its V3 model in December 2024 and R1 model in January 2025.[44] We observed significantly more references to

[41] Source documents held by Insikt Group.
[42] Source documents held by Insikt Group.
[43] Source documents held by Insikt Group.
[44] Matt O'Brien, "What is DeepSeek, the Chinese AI company upending the stock market?," *AP*, January 27, 2025, *https://apnews.com/article/deepseek-ai-china-f4908eaca221d601e31e7e3368778030*; Cade Metz, "What to Know About DeepSeek and How It Is Upending A.I.," *The New York Times*, January 27, 2025, *https://www.nytimes.com/2025/01/27/technology/what-is-deepseek-china-ai.html*.

DeepSeek in PLA and Chinese defense industry procurement records than other Chinese models like Zhipu AI and Tsinghua University's ChatGLM and Alibaba Cloud's Qwen.[45]

Generative AI Military Intelligence Providers

This section profiles two companies that likely provide generative AI technology to the PLA for intelligence purposes: DataExa (渊亭科技; 厦门渊亭信息科技有限公司; Xiamen Yuanting Information Science and Technology Co. Ltd.) and TRS (拓尔思信息技术股份有限公司; TRS Information Technology Co. Ltd.). In addition to developing generative AI-based military intelligence products, both companies have very likely functioned as OSINT platform providers for the PLA. DataExa claims to have used domestic and foreign LLMs to develop a military large model with intelligence applications. TRS claims to have used models from DeepSeek to develop an OSINT large model and a strategic simulation large model that integrates OSINT and has military applications; the company has reportedly provided its OSINT large model to the PLA.

DataExa

***Figure 3**: DataExa's logo (Source: DataExa[46])*

DataExa was founded in 2014 and claims to specialize in technologies like large models, machine learning, deep learning, and reinforcement learning.[47] National defense is among DataExa's main areas of business, and the company's clients reportedly include the Central Military Commission (CMC) Science and Technology Commission, the CMC Equipment Development Department, PLA theater commands, the former PLA Strategic Support Force, the PLA Navy, the PLA Army, China Electronics Technology Group Corporation (CETC; 中国电子科技集团公司), NORINCO Group, and national defense research institutes.[48] The company is an OSINT platform provider for the PLA; Insikt Group previously profiled DataExa in our 2023 report on the PLA's embrace of OSINT.[49]

According to one of DataExa's websites, the company offers a "Tianji military large model" (天机军事大模型), which DataExa describes as China's first military large model.[50] The capabilities of DataExa's model reportedly include "information acquisition," "understanding analysis," "knowledge reasoning," "plan generation," "plan optimization," "modal perception," and "multi-intelligent agent." The company claims its military large model serves over 70 military application scenarios, including combat

[45] Source documents held by Insikt Group.
[46] "About DataExa" [关于渊亭], DataExa [渊亭科技受], *https://web.archive.org/web/20250406061509/https://www.utenet.com/about*.
[47] Ibid.
[48] Haver, "Private Eyes"; "About DataExa."
[49] Haver, "Private Eyes"; source documents held by Insikt Group.
[50] "Tianji Military Large Model" [天机军事大模型], DataExa Defense [渊亭防务], *https://web.archive.org/web/20240617060902/https://www.utenet.com/military-model*.

command, unmanned system swarm coordination, and strategy simulation. DataExa provides detailed descriptions of several application scenarios, including "military intelligence analysis." According to the company, its Tianji military large model combines efficient natural language processing with multiple rounds of training based on years of military intelligence big data. The model can reportedly clean, extract, and organize military intelligence data and generate reports, improving the efficiency of analysis compared to traditional, manual methods.

In November 2023, DataExa published a press release and promotional video for the Tianji military large model, offering additional insight into the model and its intelligence applications.[51] The press release and video claim that the Tianji military large model draws information from DataExa's various intelligence platforms, that the model was designed to process and analyze massive amounts of OSINT data, that it can cross-verify the authenticity and effectiveness of intelligence, and that the model can train chatbots to engage in real-time question answering with intelligence analysts. The video further suggests that the Tianji military large model can use OSINT data to identify potential threats and trends, perform monitoring and early warning functions, and provide timely and feasible decision-making recommendations.

DataExa likely used multiple domestic and foreign generative AI models to develop the Tianji military large model. In an October 2023 press release, DataExa states that it used its LLM operations platform to develop the Tianji military large model and that this platform integrates multiple models, including Zhipu AI and Tsinghua University's ChatGLM, Meta's Llama, and BigScience's BLOOM.[52] Video footage of the Tianji military large model user interface published in November 2023 indicates that DataExa's model uses ChatGLM for its intelligence chatbot. In September 2024, DataExa filed a patent application for a military large model fine-tuning method that uses Alibaba Cloud's Qwen as a pre-trained model.[53] The proposed method aims to improve the accuracy and reliability of the military large model, reduce the probability of hallucination, and ensure the "authenticity and authority of military information."

TRS

***Figure 4**: TRS's logo (Source: TRS[54])*

[51] "DataExa Tianji Military Large Model: Intelligentized Open-Source Intelligence Analysis and Application [渊亭天机·军事大模型：智能化开源情报分析与应用], DataExa Defense [渊亭防务], November 21, 2023, *https://web.archive.org/web/20250421194238/https://www.utenet.cn/news-details?id=71*.
[52] "Helping the Birth of Large Models, Uncovering DataExa LLMOps Capability Borders" [助力大模型落地，起底渊亭LLMOps能力边界], DataExa [渊亭科技], *https://web.archive.org/web/20250122015819/https://www.utenet.com/news-details?id=63*.
[53] Hong Wanfu [洪万福] et al., "A fine-tuning training method, device and equipment for a large military model," Google Patents, September 12, 2024, *https://patents.google.com/patent/CN118780362A/en*.
[54] "Company Introduction" [公司简介], TRS [拓尔思], *https://web.archive.org/web/20241226154342/https://www.trs.com.cn/ljwm/gsjj/*.

TRS was founded in 1993 and specializes in big data, AI, and data security, according to its website.[55] Among other products and services,[56] the company offers an OSINT platform with national defense applications, which it has displayed at a major military-civil fusion exhibition hosted by the CMC Equipment Development Department.[57] In its 2024 annual report, TRS states that its OSINT business generated 25 million RMB ($3.4 million) in revenue in 2024 and added 90 new defense customers.[58] The company claims to have expanded its investment in global OSINT data collection and governance in 2024, surpassing 500 billion pieces of collected data, covering 196 countries and regions, supporting over 130 languages, and generating over 13 million RMB ($1.8 million) in revenue from data services related to protecting China's overseas interests. To further grow its OSINT business, TRS reportedly optimized its intelligence analysis technologies, integrated emerging technologies like AI and machine learning, and improved decision-making support. TRS further claims to have made efforts to expand its overseas business in the Middle East, Southeast Asia, Latin America, and Belt and Road Initiative countries by offering OSINT services through a subsidiary in Hong Kong.

TRS launched a generative AI model known as the "Tuotian large model" (拓天大模型) in June 2023 and, according to its 2024 annual report, adopted a DeepSeek model to serve as a foundation model for the Tuotian large model in April 2024.[59] This change, combined with an AI agent mechanism, reportedly produced a major increase in the Tuotian large model's "practical effectiveness." The report states that TRS applied its Tuotian large model and AI agent to over 40 projects across nine major domains, producing nine domain-specialized large models derived from the Tuotian large model, including an OSINT large model. The applications of the OSINT large model reportedly include "OSINT searches, intelligent question answering, intelligence judgement and analysis, target tracking and recommendation, and report compilation." The 2024 annual report suggests that customers of the OSINT large model include "multiple defense departments," PLA National Defense University (中国人民解放军国防大学), and NUDT, among others.

A March 2025 press release from TRS reveals that the company later integrated DeepSeek's R1 model into the Tuotian large model.[60] TRS appears to have used DeepSeek R1 to develop a strategic simulation large model (derived from the Tuotian large model) that integrates OSINT and is supposedly capable of supporting real-time event analysis, wartime simulations, long-term planning, strategic research, crisis response, and high-level decision-making in the military domain. The press release highlights several potential OSINT-driven applications of this simulation model, including "predicting international crises

[55] Ibid.
[56] Devin Thorne, "Evaluating the Utility of Global Data Collection by Chinese Firms for Targeted Propaganda," *China Brief* 20, no. 19 (2020), *https://jamestown.org/program/evaluating-the-utility-of-global-data-collection-by-chinese-firms-for-targeted-propaganda/*.
[57] "Tianmu Open-Source Intelligence Service Platform" [天目开源情报服务平台], TRS [拓尔思], *https://web.archive.org/web/20241225222438/https://www.trs.com.cn/cpfw/dsj/sjyy/tm_kyqb_/*; "TRS Appears at Third Military Civil Fusion Result Exhibition, and Displays Latests Products [拓尔思亮相第三届军民融合成果展，并展示最新产品], TRS [拓尔思] , September 26, 2017, *https://web.archive.org/web/20250502223825/https://www.trs-dsj.com.cn/News/detail_company/5_2017-09-26.html*.
[58] "TRS Information Technology Co. Ltd. 2024 Annual Report Summary" [拓尔思信息技术股份有限公司 2024 年年度报告摘要], QQ [腾讯网], April 18, 2025, *https://web.archive.org/web/20250425153458/https://file.finance.qq.com/finance/hs/pdf/2025/04/18/1223130051.PDF*.
[59] "TRS Tuotian Large Model Official Released, Focus on Large Model Scenario Applications and Reaching the Ground in Industry" [拓尔思拓天大模型正式发布，聚焦大模型场景化应用和行业落地], TRS [拓尔思], June 29, 2023, *https://web.archive.org/web/20240719031219/https://www.trs.com.cn/ljwm/xwhd/202306/t20230629_10053.html*; "TRS Information Technology Co. Ltd. 2024 Annual Report Summary."
[60] "Tuotian DeepSeek Version x Event Reasoning Knowledge Graph: Intelligent Game Decision-Making Paradigm Breakthrough" [拓天DeepSeek版×事理图谱：智能博弈决策范式新突破], TRS [拓尔思], March 11, 2025, *https://web.archive.org/web/20250425153139/https://www.trs.com.cn/ljwm/xwhd/202503/t20250311_10514.html*.

escalation paths," "evaluating national intervention chain reactions," "joint operations planning and simulation," and "strategic deterrence capabilities assessment."

Outlook

Challenges for the PLA

Despite the apparent enthusiasm for generative AI expressed by some elements of the PLA, the extent to which the PLA will integrate this technology into the intelligence cycle — and the ultimate effectiveness of this integration — remains unclear. Numerous plausible scenarios could result from the PLA's application of generative AI to intelligence work, including significant improvements, modest improvements, missed improvements, avoided failures, modest failures, and significant failures, which are detailed in **Table 7**. The core variables in these scenarios are the outcomes of the PLA's experiments with generative AI, the PLA's ability to accurately assess these outcomes, and the PLA's ability to appropriately apply generative AI to intelligence work based on these outcomes and assessments. Improvements to the PLA's intelligence work would likely include increases to the speed, efficiency, accuracy, and scale of intelligence tasks as well as cost reductions, and would likely enable better decision-making. Intelligence failures would likely involve decreases in the accuracy of intelligence that degrade the quality of decision-making.

Scenario	Description
Significant Improvements	The PLA's experiments with generative AI produce overwhelmingly positive outcomes and the PLA accurately assesses these outcomes, leading the PLA to apply generative AI to a wide range of intelligence tasks, resulting in significant improvements to many aspects of the PLA's intelligence work.
Modest Improvements	The PLA's experiments with generative AI produce a mix of positive and negative outcomes, and the PLA accurately assesses these outcomes, leading the PLA to cautiously use generative AI for a limited number of intelligence tasks that generative AI is capable of handling, resulting in modest improvements to the PLA's intelligence work.
Missed Improvements	The PLA experiments with generative AI produce a mix of positive and negative outcomes, but the PLA decides that generative AI is too risky, leading the PLA to largely abandon the intelligence applications of generative AI, resulting in missed opportunities for improvements to the PLA's intelligence work.
Avoided Failures	The PLA's experiments with generative AI produce overwhelmingly negative outcomes, and the PLA accurately assesses these outcomes, leading the PLA to largely abandon the intelligence applications of generative AI, resulting in the PLA avoiding intelligence failures.

Modest Failures	The PLA's experiments with generative AI produce a mix of positive and negative outcomes, but the PLA misinterprets some of the negative outcomes, leading the PLA to inappropriately apply generative AI to certain intelligence tasks, resulting in modest intelligence failures for the PLA.
Significant Failures	The PLA's experiments with generative AI produce largely negative negative outcomes, but the PLA incorrectly assesses these outcomes to be overwhelmingly positive, leading the PLA to incautiously apply generative AI to a wide range of intelligence tasks, resulting in significant intelligence failures for the PLA.

***Table 7**: Plausible scenarios resulting from the PLA's application of generative AI to intelligence work (Source: Insikt Group)*

Moreover, pressure to conform to the CCP's ideological requirements can reportedly diminish the objectivity of intelligence analysis within China's intelligence services, and the PLA's use of generative AI for intelligence analysis risks exacerbating these ideological distortions.[61] Not only will a generative AI model trained on ideologically biased intelligence analysis very likely generate ideologically biased intelligence analysis, generative AI developers in the PLA and China's defense industry will also likely prioritize the development of ideologically compliant models to increase the chances of the PLA adopting their models. If the PLA adopts ideologically biased generative AI models, these models could reinforce the reported tendency of China's intelligence services to assume the worst of foreign actors, potentially fueling miscalculations, overreactions, and escalatory behavior.

Challenges for the West

The PLA and Chinese defense industry's use of foreign open-source LLMs for intelligence purposes illuminates two technology transfer challenges for the West. The first is the risk of individuals and organizations involved in developing open-source generative AI models unintentionally supporting China's military, possibly in violation of ethical guidelines, company policies, or government regulations. The second is the difficulty of using export controls to prevent military end users in China from accessing open-source generative AI models.[62] Proprietary generative AI models are also vulnerable to technology transfer risks, as demonstrated by the aforementioned examples of the PLA and China's defense industry using (or discussing the use of) OpenAI's models for intelligence purposes and reports of DeepSeek allegedly distilling OpenAI's models to support the development of its own models, which the PLA is likely using for intelligence tasks.[63] There likely are no simple solutions to these challenges,

[61] Peter Mattis and Matthew Brazil, *Chinese Communist Espionage: An Intelligence Primer* (Naval Institute Press, 2019); Matthew Brazil, "China: The Fearful Intelligence Culture" in *The Handbook of Asian Intelligence Cultures*, ed. Ryan Schaffer (Rowman & Littlefield, 2022).
[62] Claudia Wilson and Emmie Hine, "Export Controls on Open-Source Models Will Not Win the AI Race," *Just Security*, February 25, 2025, *https://www.justsecurity.org/108144/blanket-bans-software-exports-not-solution-ai-arms-race/*; Emily S. Weinstein and Kevin Wolf, "For Export Controls on AI, Don't Forget the 'Catch-All' Basics," Center for Security and Emerging Technology, July 5, 2023, *https://cset.georgetown.edu/article/dont-forget-the-catch-all-basics-ai-export-controls/*; Michael C. Horowitz, "What to Know About the New U.S. AI Diffusion Policy and Export Controls," Council on Foreign Relations, January 13, 2025, *https://www.cfr.org/blog/what-know-about-new-us-ai-diffusion-policy-and-export-controls*.
[63] Cristina Criddle and Eleanor Olcott, "OpenAI says it has evidence China's DeepSeek used its model to train competitor," *Financial Times*, January 28, 2025, *https://www.ft.com/content/a0dfedd1-5255-4fa9-8ccc-1fe01de87ea6*; Dina Bass and Shirin Ghaffary, "Microsoft Probing If DeepSeek-Linked Group Improperly Obtained OpenAI Data," *Bloomberg*, January 29, 2025, *https://www.bloomberg.com/news/articles/2025-01-29/microsoft-probing-if-deepseek-linked-group-improperly-obtained-openai-data*.

and addressing these challenges will likely require coordination between private industry and governments.

Intelligence agencies in the West and elsewhere should further prepare for the possibility of Chinese counterintelligence organizations using generative AI to produce inauthentic but convincing information to mislead intelligence personnel and degrade the intelligence value of open-source information. Insikt Group makes this recommendation based on the following evidence:

- Elements of the PLA and China's public security system have demonstrated concern that foreign counterintelligence organizations could use generative AI in a similar manner against China.[64]
- The PLA very likely views foreign OSINT activity as a threat, and China's party-state system has already taken significant steps to counter foreign OSINT collection in recent years.[65]
- The PLA has demonstrated interest in using generative AI to support cyber-enabled influence operations;[66] OSRAC filed a patent application in April 2024 — which was granted in April 2025 — that proposes using OpenAI's Sora video generation model to facilitate cognitive warfare and suggests that AI-generated content can effectively influence adversary decision-making;[67] and Chinese covert influence networks have used generative AI to carry out online influence operations.[68]

Government, military, and intelligence organizations with an interest in China, as well as generative AI developers in the private sector and academia, should continue to monitor PLA and Chinese defense industry sources to track the PLA's efforts to apply generative AI to intelligence work, assess the effectiveness of the PLA's generative AI-based intelligence capabilities, evaluate technology transfer and counterintelligence risks, and develop appropriate countermeasures.

[64] Yang, "Be on Alert for Deepfake Technology"; Tang et al., "Empowering National Defense Science and Technology Intelligence"; Li and Wu, "The Impact of False Information on Open-Source Intelligence."
[65] Haver, "Private Eyes"; source documents held by Insikt Group.
[66] Nathan Beauchamp-Mustafaga, "Exploring the Implications of Generative AI for Chinese Military Cyber-Enabled Influence Operations," RAND, February 1, 2024, *https://www.rand.org/pubs/testimonies/CTA3191-1.html*; Nathan Beauchamp-Mustafaga et al., "Dr. Li Bicheng, or How China Learned to Stop Worrying and Love Social Media Manipulation," October 1, 2024, *https://www.rand.org/pubs/research_reports/RRA2679-1.html*.
[67] Zhao Yuenchen [**赵悦辰**] et al., "Construction method and application of cognitive war system based on aragonic video," Google Patents, April 17, 2024, *https://patents.google.com/patent/CN118568297A/en*; source documents held by Insikt Group.
[68] "Disrupting deceptive uses of AI by covert influence operations," OpenAI, May 30, 2024, *https://openai.com/index/disrupting-deceptive-uses-of-AI-by-covert-influence-operations/*; Clint Watts, "China tests US voter fault lines and ramps AI content to boost its geopolitical interests," Microsoft, April 4, 2024, *https://blogs.microsoft.com/on-the-issues/2024/04/04/china-ai-influence-elections-mtac-cybersecurity/*; David E. Sanger and Steven Lee Myers, "Those Hawaii Wildfires? Caused by U.S. Weather Weapon, China Says," *The New York Times*, September 11, 2023, *https://www.nytimes.com/2023/09/11/us/politics/china-disinformation-ai.html*.

Appendix A: Glossary of Generative AI Terminology

This appendix provides basic definitions for the generative AI terminology used in this report. It is intended for readers who are not familiar with generative AI.

Term	Definition
Agent	AI agents are software systems that can reason, act, observe, plan, collaborate, and self-refine to complete tasks on behalf of users.[69]
Corpus (Corpora)	Corpora are the datasets used to train generative AI models.[70]
Fine-Tuning	Fine-tuning a model involves providing further training to a pre-trained model to help the model perform specialized tasks more effectively.[71]
Foundation Model	Foundation models are generally trained on large, unlabeled datasets, are capable of carrying out a variety of tasks, and can be tuned for more specific applications.[72]
Generative AI	Generative AI models can generate original text, images, video, and other content based on a large training dataset.[73]
Instruction Tuning	Instruction tuning is a form of fine-tuning that involves training a model on instruction prompts to improve its ability to follow instructions.[74]
Large Language Model (LLM)	LLMs are models trained on large text datasets that can engage in text-related tasks like text generation, text translation, text summarization, and question answering; they are considered to be a form of generative AI but are not limited to generative tasks.[75]

[69] "What is an AI agent?," Google Cloud, *https://cloud.google.com/discover/what-are-ai-agents?hl=en*; "Machine Learning Glossary: Generative AI," Google for Developers, *https://developers.google.com/machine-learning/glossary/generative*.
[70] "GPT-4," OpenAI, March 14, 2023, *https://openai.com/index/gpt-4-research/*; Pierre-Carl Langlais, "Releasing Common Corpus: the largest public domain dataset for training LLMs," March 20, 2024, *https://huggingface.co/blog/Pclanglais/common-corpus*.
[71] "Fine-tuning," OpenAI Platform, *https://platform.openai.com/docs/guides/fine-tuning*; "Machine Learning Glossary: Generative AI"; "Methods for adapting large language models," Meta, August 7, 2024, *https://ai.meta.com/blog/adapting-large-language-models-llms/*.
[72] "What are foundation models?," IBM, *https://research.ibm.com/blog/what-are-foundation-models*; Rick Merritt, "What Are Foundation Models?," Nvidia, February 11, 2025, *https://blogs.nvidia.com/blog/what-are-foundation-models/*; "Tuning foundation models," IBM, MArch 4, 2025, *https://www.ibm.com/docs/en/watsonx/saas?topic=solutions-tuning-models*; Toner, "What Are Generative AI, Large Language Models, and Foundation Models?"
[73] "Generative models," OpenAI, June 16, 2016, *https://openai.com/index/generative-models/*; "Machine Learning Glossary: Generative AI"; Toner, "What Are Generative AI, Large Language Models, and Foundation Models?"
[74] "Machine Learning Glossary: Generative AI."
[75] "Large language models" Google for Developers, *https://developers.google.com/machine-learning/crash-course/llm*; "Better language models and their implications," OpenAI, February 14, 2019, *https://openai.com/index/better-language-models/*; "Large Language Model (LLM)," Microsoft AI Tour, *https://microsoft.github.io/Workshop-Interact-with-OpenAI-models/llms/*; Toner, "What Are Generative AI, Large Language Models, and Foundation Models?"

Multimodal Model	A multimodal model can accept inputs or generate outputs from multiple different categories of data, such as text, images, video, and audio.[76]
Parameter-Efficient Fine-Tuning	Parameter-efficient fine-tuning is an efficient, low-resource, and cost-effective fine-tuning method that fine-tunes a relatively small number of parameters, which are the weights and biases used during model training.[77]
Pre-Training	Pre-training is the initial training phase of AI model development that involves training a model on a large dataset.[78]
Prompt Engineering	Prompt engineering is the process of testing, refining, and applying prompts, including by providing context, instructions, and examples, to a generative AI model to help the model generate accurate and relevant outputs.[79]

Table 8: *Generative AI terms and definitions (Source: Insikt Group)*

[76] "Machine Learning Glossary: Generative AI"; "GPT-4"; "Multimodal generative AI systems," Meta, December 12, 2023, *https://ai.meta.com/tools/system-cards/multimodal-generative-ai-systems/*.
[77] "Machine Learning Glossary: Generative AI"; "Introduction to tuning," Google Cloud, *https://cloud.google.com/vertex-ai/generative-ai/docs/models/tune-models*.
[78] "Machine Learning Glossary: Generative AI"; "Methods for adapting large language models."
[79] "Prompt engineering: overview and guide," Google Cloud, *https://cloud.google.com/discover/what-is-prompt-engineering?hl=en*; "Prompting," Meta, *https://www.llama.com/docs/how-to-guides/prompting/*; "Machine Learning Glossary: Generative AI."

Recorded Future reporting contains expressions of likelihood or probability consistent with US Intelligence Community Directive (ICD) 203: Analytic Standards (published January 2, 2015). Recorded Future reporting also uses confidence level standards employed by the US Intelligence Community to assess the quality and quantity of the source information supporting our analytic judgments.

About the Author

Zoe Haver is a senior threat intelligence analyst on Insikt Group's Global Issues team. Her research focuses on the People's Liberation Army and China-related security issues. Before joining Recorded Future, she worked on these topics for Radio Free Asia, the Center for Advanced China Research, SOSi's Center for Intelligence Research and Analysis, the US Naval War College China Maritime Studies Institute, C4ADS, and other organizations. She received her BA from George Washington University and is proficient in Mandarin Chinese.

About Insikt Group®

Recorded Future's Insikt Group, the company's threat research division, comprises analysts and security researchers with deep government, law enforcement, military, and intelligence agency experience. Their mission is to produce intelligence that reduces risk for customers, enables tangible outcomes, and prevents business disruption.

About Recorded Future®

Recorded Future is the world's largest threat intelligence company. Recorded Future's Intelligence Cloud provides end-to-end intelligence across adversaries, infrastructure, and targets. Indexing the internet across the open web, dark web, and technical sources, Recorded Future provides real-time visibility into an expanding attack surface and threat landscape, empowering customers to act with speed and confidence to reduce risk and securely drive business forward. Headquartered in Boston with offices and employees around the world, Recorded Future works with over 1,800 businesses and government organizations across more than 75 countries to provide real-time, unbiased, and actionable intelligence.

Learn more at recordedfuture.com